面孔可信度评价调节：经验迁移假说的提出与验证*

齐玥 [1,2] 秦邵天 [1,2] 王可昕 [1,2] 陈文锋 [1,2]

1 中国人民大学心理学系，北京 100872

2 中国人民大学心理学系实验室，北京 100872

摘要 面孔可信度评价对人际信任和合作行为有重要影响。以往研究发现面部特征、背景环境等附加信息能通过自下而上、自上而下两种方式调节人们对他人面孔可信度的评价， 但具体的起效和失效机制尚不清晰。针对这一问题， 本文提出了“经验迁移假说”， 并对两类调节方式的认知机制、经验迁移的边界条件和调节过程进行探究。本研究为深化理解人际信任的调控机制提供理论和实证依据。

关键词 信任，调节效应，面孔可信度，印象形成

分类号 B849:C91

人们常说“相由心生”，当看到一张陌生面孔时，基于面孔的视觉线索，我们会自动形成对他的评价(Olivola et al, 2014)。大量研究证实，仅通过几百甚至几十毫秒对面孔相片的观察，人们就能够对一个人的性格有所判断，并影响后续的决策和行为(Gheorghiu et al., 2017; Todorov, 2005; Todorov et al., 2015; 李庆功 等, 2020; 刘晨 等, 2019)。其中，对于面孔可信度的评价是人们对于他人的效价判断，也是对后续行为最有影响力的判断，对建立人际信任十分重要(Bonnefon et al., 2017; Jones, 2021; Shen et al., 2020; Zebrowitz & Montepare, 2005, 2008; 马凤玲 等, 2014)。

面孔可信度(facial trustworthiness)是指由外貌特征决定的个体值得信任的程度。 当遇到陌生人或陌生群体时，个体会本能地想知道陌生他人的行为意图和能力状况，这决定着个体选择趋近还是回避行为(Todorov et al., 2015)。这种由面部特征得出的内隐特质推断影响着人们的日常生活。例如，高面孔可信度的求助者更有可能得到他人的信任，进而得到帮助(王秀娟 等, 2018)。在经济管理领域，研究发现，面孔可信度与管理者的地位显著正相关，即下级人数越多，上级人数越少，可感知的可信度越高(Linke et al., 2016)；高面孔可信度的企业家

收稿日期：2021-8-27

* 国家自然科学基金(32000771)，中国人民大学科学研究基金（中央高校基本科研业务费专项资金资助）项目成果(21XNLG13)，北京市自然科学基金(5184035), 2018 年度中央高校建设世界一流大学（学科）和特色发展引导专项资金(RUCPSY0007)。

通信作者：齐玥，E-mail: qiy@ruc.edu.cn

在众筹资金时，有更高的募集成功率且能募集到更多的资金(Duan et al., 2020)；高面孔可信度的首席财务官在审计时将被审计员收取更少的审计费(Hsieh et al., 2020)。 在刑法判决中，人们更倾向于认为长相不值得信任的被告有罪或是对其处以更长时间的判罚(Jaegera et al., 2020; Wilson & Rule, 2015)。在机器人领域，面孔可信度能用于改善社会机器人与人的交流和互动(Song, 2020; Song & Luximon, 2020)。综上，面孔可信度在人们生活中有着重要且广泛的影响。

1 面孔可信度评价的国内外研究现状

目前，研究者常用的面孔刺激材料多来自于标准化的面孔图片库[†]。基于是否改变面孔相片刺激[‡]， 以往对面孔可信度评价调节的研究可分为两类， 自下而上面部特征和自上而下环境线索对面孔可信度评价的影响。

1.1 自下而上因素对面孔可信度评价的影响

个体面部特征对面孔可信度的影响，一直受到学者们的广泛关注。以往的研究发现，面孔可信度评价可以通过显著的视觉线索自下而上地进行调节，相片中面部特征(Kleisner et al., 2013; Song, 2020)、面部的典型性(Sofer et al., 2015)、面孔的宽高比(Bryan et al., 2012; Hehman et al., 2015)、面部表情(Slepian & Carr, 2019; Sutherland et al., 2017)、眼睛注视方向(Kaisler & Leder, 2016; Manssuer et al., 2015; 李永娜，昂亚轩, 2019)和头部姿势(Kaisler & Leder, 2017; Zhang et al., 2020)等差异均能影响面孔可信度的评价。例如眉毛内角上扬、颧骨突出、下巴宽平、鼻骨浅平、面部较宽、脸部皮肤光滑的人看上去更值得人信任(Bakmazian, 2014; Oosterhof & Todorov, 2008; Tsankova & Kappas, 2016)。同时，面孔的典型性程度也是影响可信度评价的重要因素,看上去越符合一般性的典型面孔，人们所知觉到的可信程度就越高(Sofer et al., 2015)。面部宽高比(fWHR)是研究者早期发现的自下而上的调节因素之一(Geniole et al., 2014; Haselhuhn & Wong, 2012; Stirrat & Perrett, 2010)。虽然 fWHR 与大部分实际的行为倾向并没有显著关联，但人们依然会利用这个线索来评估他人的反社会倾向(Kosinski, 2017; Wang et al., 2019)。有研究表明，对同一张面孔，被试倾向于判断面孔宽高比较小的面孔是相对不可信的(Costa et al., 2017)。给面部添加纹身，同样也能影响面孔评价，无

† 西方面孔数据库如芝加哥面孔数据库(The Chicago Face Database, Ma et al., 2015)、10k US Adult Faces Database (Bainbridge et al., 2013)以及中国面孔数据库如 CAS-PEAL 面孔库 (Gao et al., 2008; 张晓华 等, 2005)、CAFPS (龚栩 等, 2011)。

‡ 常用的面孔相片处理工具包括：FantaMorph、 FaceResearch、 Morpheus Photo Morpher、 Star GAN、InferFace、 MakeHuman、 FaceGen Modeller 等

论面孔自身的可信程度如何，为面孔相片加上带刺铁丝的纹身图案均提高了该个体在模拟法庭场景下被判定有罪以及判断后续再次犯罪可能性的比例(Funk & Todorov, 2013)，即面部纹身降低了面孔可信度评价。此外，面孔表情也与面孔可信度评价有关。研究发现，可信面孔更容易被感知到快乐情绪，不可信面孔更容易被感知到愤怒情绪(Oosterhof & Todorov, 2009)；同时，快乐情绪的面孔更容易被知觉为信任，愤怒情绪的面孔更容易被知觉为不可信(Caulfield et al., 2014; Sutherland et al., 2017)。面孔相片中眼睛的注视方向同样能影响个体对面孔可信度的判断，相对于斜视的个体，目光直视的个体被认为更加可信(Kaisler & Leder, 2016)。此外，头部姿势对面孔可信度评价也有影响，研究表明当目光方向不变时，水平的头部姿势比抬头和低头姿势被认为更加可信，且抬头的姿势比低头姿势被认为更加可信(Zhang et al., 2020)。但是，改变相片中面孔的视觉特征并不一定影响面孔可信度的评价。比如,调整了眼睛的注视方向后，悲伤表情面孔的可信度评价并没有发生变化(Willis et al., 2011; 李永娜，昂亚轩, 2019)。

1.2 自上而下因素对面孔可信度评价的影响

在现实生活中，面孔很少孤立出现，通常会嵌入到环境信息中。面孔可信度评价的研究发现，即使不改变面孔本身的物理特征，面孔以外的信息(如收入信息、过往的行为、背景等)也会通过自上而下的方式来调节面孔可信度的评价。

有研究发现，个体的社会地位以及所扮演的角色等会对面孔可信度的评价造成影响。Qi等人(2018)对社会经济地位如何影响面孔可信度进行了研究，结果发现，给面孔提供月收入信息可以显著影响人们对面孔可信度的评价，高收入的面孔被认为更加可信。另外，不同国家的研究者一致发现，提供过去的合作或是道德行为相关的正性信息，能够提高面孔可信度的评价(Li ei al., 2017; Mende-Siedlecki et al., 2013; Rezlescu et al., 2012; Verosky et al., 2018; Yu et al., 2014)；面孔可信度评价也会受到面孔所处视觉背景的影响(Brambilla et al., 2018)。此外，观察者本身的特征也会影响对面孔可信度的评价。例如，有研究表明评价者自身的年龄会影响面孔可信度的判断(Neys, 2015)，面部表情对成年人的面孔可信度判断的影响也要高于儿童(Caulfield et al., 2016)。

1.3 面孔可信度评价调节中的神经机制

为面孔相片相继提供冲突的外部信息，是研究面孔评价调节过程神经机制的常用范式。例如，在 Mende-Siedlecki 等人的系列研究中，通过操纵前后呈现的行为信息在效价上的一致或不一致，来观察面孔评价调节中的神经活动(e.g., Mende-Siedlecki et al., 2013; Mende-Siedlecki & Todorov, 2016)。结果发现，背内侧前额叶(dmPFC)是面孔评价中的重要脑区(e.g.,

Baron et al., 2011; Ferrari et al., 2016; Mitchell et al., 2005)，同时，面孔评价调节的过程涉及一系列分散的脑区，包括社会知觉相关的颞上沟(superior temporal sulcus; STS)、心理理论相关的颞顶联合区/顶下小叶(temporoparietal junction/inferior parietal lobule; TPJ/IPL)、期望违背相关的背侧扣带(dorsal anterior cingulate; dACC)，以及认知控制相关的外侧前额叶皮层(lateral prefrontal cortex; lPFC)。Hughes 等人(2017)的研究采用校徽作为内/外群体的面孔身份信息，采用冲突的行为信息作为印象改变的线索。结果发现，在被试评价外群体面孔时，外侧前额叶皮质(lPFC)、背侧前扣带皮层(dACC)、 颞顶联合(TPJ)、脑岛等区域的激活减弱。这些研究为我们进一步探索脑区之间如何协同作用提供了结果上的参考，证实了 dmPFC、lPFC、dACC 等几个脑区在面孔评价调节中的重要作用，因此，有必要进一步探究经验迁移后相关脑区之间的协同活动。

此外，在行为研究中发现，从最初的正面印象转变为负面印象要更容易(Albarracin & Shavitt, 2018; Cone & Ferguson, 2015)，而不可信面孔转化为可信面孔则需要具有更强诊断性的行为线索(Shen et al., 2020)。因此，本研究预期在面孔可信度评价调节过程中，不同调节方向（提高/降低）可能存在独立的神经机制。

2 经验迁移假说的理论建构

2.1 经验对于面孔可信度评价调节的重要性

首先，面孔评价中的文化特异性(Krys et al., 2014; Sofer et al., 2017)为经验迁移假说的提出提供了间接证据。Sofer 等人(2017)采用日本和以色列的面孔相片，发现被试只对本国的典型面孔可信度评价更高；Kleider-Offutt 等人(2017)发现，具有典型黑人特征(比如宽鼻子、厚嘴唇)的个体更易被分类为黑人，这种分类反过来又使这一亚群更有可能与犯罪的黑人男性刻板印象联系在一起，被知觉者认为有更低的可信度。

其次，对于婴儿和机器人的研究发现经验能够影响面孔偏好和评价。Quinn 等人(2019)总结了有关婴儿面孔偏好的研究，并提出早期抚养经验的差异(更多本族、女性抚养)导致婴儿更偏好本族面孔和女性面孔。Xiao 等人(2018)的研究发现，婴儿直到 9 个月才表现出对本族面孔与高兴乐曲这一组合的偏好。这表明婴儿在抚养过程中不断积累了本族面孔所展现出的积极情绪，进而更加熟悉本族面孔与积极情绪的关联。此外，由于人们对于机器人没有足够的先验经验，这使得研究者可以在实验室创造不同的人机互动体验，进而探测人们对于机器人的评价是否发生改变。例如，Paetzel 和 Castellano(2019)发现人们对于机器人胜任力的评价会根据实际互动结果不断更新。

2.2 经验泛化影响了新信息对于面孔可信度评价的调节效果

经验泛化，即人们能否将经验中线索与信任倾向的关联迁移至当前线索与信任倾向的关联。关联学习领域提出，新旧线索在概念或知觉上的相似性以及线索与行为之间的关联程度(概率)能够影响学习效果(Einstein et al., 2005; Rescorla, 1976; Rescorla & Furrow, 1977; Verosky & Todorov, 2010)。FeldmanHall 等人(2018)发表在 PNAS 的研究采用关联学习范式，证实当前合作者与先前经验学习中的合作者或背叛者长相上的相似性，能够影响人们的信任行为。该研究首次证实了人们能将经验中面孔与信任倾向的关联迁移到对新面孔的信任行为中，并且证实了新旧面孔在知觉上的相似程度是影响经验能否迁移到新面孔的关键。这一结果为经验迁移假说提供了直接证据。遗憾的是，现实生活中面孔的相似性往往是亲缘关系的体现。对于长相完全陌生的面孔，线索在概念上的相似性、与信任倾向的关联程度以及不同线索间的对比是否影响经验泛化仍值得进一步研究。采用关联学习范式，在实验室建立一种崭新的经验，并将其与可信度联结，将成为一种可行的实验方法。

2.3 经验迁移假说的提出

人们对于面孔可信度的评价是快速、自动化的，这种评价并非一成不变，面孔可信度的评价可以被自下而上、自上而下两类调节因素改变。为什么这些调节因素能够影响人们对面孔可信度的评价？无论是对于面孔纹身的偏见，还是场景一致性的预期，均源自评价者有关信任的先验知识。基于此，本研究提出了“经验迁移假说”来提供一种可能的解释，即某一线索能否调节当前面孔的可信度评价受到了评价者先前经验及其迁移结果的影响。“经验迁移假说”包含两层含义: (1)经验是指人们学习到的线索与信任倾向之间的联结关系，当人们不具备相关经验时，线索将失效; (2)经验泛化决定了经验迁移是否成功，当新线索与经验中线索的相似度较低，或先前线索与信任倾向的关联不强时，人们都难以将新线索与信任倾向关联，经验无法泛化，调节因素就会失效。因此，当出现两种冲突的调节线索时，与先前线索相似性更高、与信任倾向关联更紧密的线索，将主导调节效果。

3 问题提出

在复杂的社会情境中，个体获得信息的渠道是多样的， 因此现实中人们对他人面孔可信度的评价可能会发生改变。发表在 Current Opinion in Psychology 的综述明确提出之前的研究大多关注“静态”的面孔评价，而对面孔评价调节机制的研究远远不够(Mende-Siedlecki, 2018)。随着网络视频等基于面孔交互场景的增加，面孔可信度作为面孔评价中的核心(Wilson & Rule, 2017; 朱梦音 等, 2014)，其评价会受到哪些因素调节，以及如何受到这些因

素调节将成为未来几年这一领域的研究趋势。

有研究表明，面孔可信度评价会受到相片变化和额外信息的调控(e.g., Bryan et al., 2012; Fareri et al., 2012; Mende-Siedlecki et al., 2013; Strachan et al., 2017; Yu et al., 2014)。例如，当可信面孔处于威胁背景时，背景本身的负性效价会降低被试对面孔可信度的判断(Brambilla et al., 2018)；给面孔添加月收入信息能够影响面孔可信度评价，但添加国籍信息不会影响(Qi et al., 2018)。然而，无论是单一因素还是多种因素共同作用，针对面孔可信度评价的调节均可能失效。

面孔可信度评价的调节因素为什么会失效？现有的理论假设尚不能解释面孔评价中的调节机制。目前主流的情绪泛化理论(Oosterhof & Todorov, 2008, 2009; Zebrowitz et al., 2003)认为，可信面孔与快乐表情、不可信面孔与愤怒表情的相似影响了人们的评价。该理论从面孔刺激本身出发，自下而上地解释了情绪线索的相似性对于面孔评价的影响，但是无法解释面孔特征不变时，自上而下的不同身份背景所导致的可信度评价差异。信息约束假说则是从自上而下的视角，提出环境信息在社会认知加工的早期阶段限制了部分感知信息加工(Brambilla et al., 2018)。但是，这无法解释为什么环境信息在有些情况下不起作用。综上，以往研究已经提出了自上而下和自下而上两类面孔可信度评价的调节因素，但是对于二者在调节作用上起效和失效的机制尚不明确。基于以往的研究，本研究围绕面孔可信度评价的调节提出了“经验迁移假说”。在经验迁移假说框架下，本研究提出以下三个问题:

问题 1: 经验中线索与信任倾向的关联是否能通过自下而上以及自上而下两种途径调节面孔可信度的评价？

问题 2: 从调节效果上，如果经验中线索与信任倾向的关联能够调节新面孔的可信度评价，那么经验泛化如何影响调节效果？

问题 3: 从调节过程上，提高/降低面孔可信度评价在神经机制上有哪些异同？

4 研究构想

以往研究发现了面孔可信度评价调节失效的现象，但尚未明确调节失效的原因。因此，基于申请人前期的研究发现，本文将围绕面孔可信度评价调节中的“经验迁移假说”，从自下而上以及自上而下两种调节路径出发，一方面通过行为实验验证经验在面孔可信度调节中的重要作用，另一方面通过脑成像技术，对面孔可信度调节的认知神经机制展开研究。

4.1 面部线索与环境线索对面孔可信度的影响

验证“经验迁移假说”，即经验中线索与信任倾向的关联能够通过自下而上以及自上而下两种方式调节新面孔的可信度评价。采用关联学习范式，在经验学习阶段建立线索与信任倾向间的关联。然后，让被试评价具备相关线索的新面孔，从而研究先前经验中线索与信任/不信任的关联能否通过自下而上(面部线索)、自上而下(环境线索)两种方式对面孔可信度评价进行调节。

4.2 线索相似程度、关联程度及线索间对比的影响

从调节效果上，探索经验泛化对于面孔可信度评价调节的影响。在经验迁移过程中，人们能否将经验中的线索与信任倾向的关联泛化可能是影响面孔可信度评价调节效果的重要因素。影响经验泛化的因素包含：经验中线索与新线索的相似程度；经验中线索与信任倾向的关联程度；呈现多种线索时，线索之间的对比。因此，研究二将从这三个不同的角度，进一步研究经验泛化如何导致面孔可信度评价的不同调节效果，以此探索经验迁移的边界条件。

4.3 面孔可信度调节方向的影响

从调节过程上，探索面孔可信度评价调节的神经机制及不同调节方向的影响。不同的调节方向可能有着不同的面孔可信度调节过程。行为研究能够解决先前经验是否以及多大程度上影响面孔可信度评价的调节，采用脑成像技术，则可以进一步分析不同的调节方向上相关脑区间的功能交互。结合 PPI 和 DCM 分析方法可推测关键脑区之间的因果结构，阐明面孔可信度评价调节中不同调节方向有哪些共享以及特异的神经机制。

5 总结

本研究从面孔这一人际交往中的重要线索出发，对面孔可信度评价的调节机制进行深入探讨， 在理论上提出了“经验迁移假说”， 深化了有关信任行为发生机制的探索； 在实践上也将指导人们更有效地调整自身的信任水平。

在理论上，首先，先前的研究设计难以对被试记忆中的先验知识和新线索进行直接比对，因此，本研究引入“经验”这一变量，创造性地在实验室条件下模拟两类调节因素的起效过程，为之后引导和改善人际信任提供可借鉴的方法和实验依据。其次，本研究将系统性回答人们如何在给定的第一印象上利用经验调节面孔可信度的评价，探索经验迁移的边界条件以及调节方向对面孔可信度评价的影响，找出调节失效的原因，了解人际信任的调节机制。以往研究发现 56~67%以上的面部合成比例能够影响人们的合作行为(FeldmanHall et al., 2018)，从知觉上的线索相似性角度给出了经验迁移的边界条件。但这种大幅改变面孔特征的调节方

式在现实中的可推广性较差。基于关联学习领域的研究发现，本研究将从概念上的线索相似性、线索与信任倾向的关联程度以及不同线索间的对比三个不同角度,进一步探索经验迁移的边界条件。最后，在网络相片、视频等新的人际交互模式盛行的当下，以往研究报告了想象信任/不信任在脑影像证据上的不对称(Filkowski et al., 2015)，但在面孔可信度评价的调节过程中，相关脑区之间的协同活动是否存在提高/降低的不对称效应，目前尚缺少相关研究。以背内侧前额(dmPFC)、外侧前额叶皮质(lPFC)、背侧前扣带皮层(dACC)、颞顶联合区(TPJ)为代表的相关脑区，是影响面孔评价和调节的重要脑区。通过实验模拟经验产生，本研究将探索基于经验迁移的面孔可信度评价调节所依赖的神经机制，以及提高/降低评价的过程有哪些神经机制是共享的，哪些是各自独立的？上述问题的解决将有助于深入理解面孔可信度评价的调节机制，能够帮助人们在保持原有面容的前提下，通过建立外部经验改变自身的可信程度。本研究将为进一步预测以及调整人际信任行为提供依据，为更好地营造和谐可信的人际关系和社会氛围提供建议。本研究结果也将有助于社交机器人的设计，为构建更好的人机信任提供参考(Zhang et al., 2020)。本研究结果对于深化理解人际信任的调节机制，有着重要的理论意义。

在潜在应用领域，目前人们越来越多的采用图片、视频的方式进行社会互动，在无法改变长相及其所带来的信任偏向时，本研究的研究成果将有助于建立成功的经验迁移，提高面孔可信度评价(比如通过改变录像背景提高讲话人的可信度)，从而改善人际信任。

参考文献

龚栩，黄宇霞，王妍，罗跃嘉. (2011). 中国面孔表情图片系统的修订. *中国心理卫生杂志*, *25*(1), 40–46.

李庆功，王震炎，孙捷元，师妍. (2020). 网约车场景中声誉和面孔可信度对女性信任判断的影响以及直觉性思维的调节作用. *心理科学进展*, *28*(5), 746–751.

李永娜，昂亚轩. (2019). 面孔性别，注视方向和面部表情对面部可信度判断的影响. *心理学进展*, *9*(3), 609–617.

刘晨，温芳芳，佐斌. (2019). 以貌取人可行吗？——基于面孔的特质推理及准确性. *心理科学*, *43*(1), 150–156.

马凤玲，郑婷婷，蔡维香，徐芬. (2014). 信任倾向和面孔可信度对小学儿童信任判断的影响. *应用心理学*, *20*(3), 216–226.

王秀娟，王娜，韩尚锋，刘燊，张林. (2018). 面孔可信度对助人行为的影响: 依恋安全的调节作用. *心理学报*, *50*(11), 1292–1306.

张晓华，山世光，曹波，高文，周德龙，赵德斌. (2005). 大规模中国人脸图像数据库及其基本评测介绍. *计算机辅助设计与图形学学报*, *17*(1), 9–17.

朱梦音，江琦，侯敏，陈潇. (2014). 面孔可信任度评价的研究回顾. *教育生物学杂志*, *2*(3), 189–194.

Albarracin, D., & Shavitt, S. (2018). *Attitudes and attitude change*. *Annual Review of Psychology*, *69*, 299–327.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of*

Experimental Psychology: General, *142*(4), 1323–1334.

Bakmazian, A. (2014). The man behind the beard: Perception of men's trustworthiness as a function of facial hair. *Psychology*, *05*(3), 1185–1191.

Baron, S. G., Gobbini, M. I., Engell, A. D., & Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social Cognitive and Affective Neuroscience, 6*(5), 572–581.

Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2017). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science*, *26*(3), 276–281.

Brambilla, M., Biella, M., & Freeman, J. B. (2018). The influence of visual context on the evaluation of facial trustworthiness. *Journal of Experimental Social Psychology*, *78*, 34–42.

Bryan, R., Perona, P., & Adolphs, R. (2012). Perspective distortion from interpersonal distance is an implicit visual cue for social judgments of faces. *PLoS ONE*, *7*(9), e45301.

Caulfield, F., Ewing, L., Bank, S., & Rhodes, G. (2016). Judging trustworthiness from faces: Emotion cues modulate trustworthiness judgments in young children. *British Journal of Psychology*, *107*(3), 503–518.

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial trustworthiness judgments in children with ASD are modulated by happy and angry emotional cues. *PLoS ONE*, *9*(5), e97644.

Cone, J., & Ferguson, M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology*, *108*(1), 37–57.

Costa, M., Lio, G., Gomez, A., & Sirigu, A. (2017). How components of facial width to height ratio differently contribute to the perception of social traits. *PLoS ONE*, *12*(2), e0172739.

Duan, Y., Hsieh, T.-S., Wang, R. R., & Wang, Z. (2020). Entrepreneurs' facial trustworthiness, gender, and crowdfunding success. *Journal of Corporate Finance*, *64*, 101693.

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology: General*, *134*(3), 327–342.

Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in Neuroscience*, *6*, 148.

FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). Stimulus generalization as a mechanism for learning to trust. *Proceedings of the National Academy of Sciences*, *115*(7), E1690–E1697.

Ferrari, C., Lega, C., Vernice, M., Tamietto, M., Mende-Siedlecki, P., Vecchi, T., … Cattaneo, Z. (2016). The dorsomedial prefrontal cortex plays a causal role in integrating social impressions from faces and verbal descriptions. *Cerebral Cortex*, *26*(1), 156–165.

Filkowski, M. M., Anderson, I. W., & Haas, B. W. (2015). Trying to trust: Brain activity during interpersonal social attitude change. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(2), 325–338.

Funk, F., & Todorov, A. (2013). Criminal stereotypes in the courtroom: Facial tattoos affect guilt and punishment differently. *Psychology, Public Policy, and Law*, *19*(4), 466–478.

Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D. (2008). The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *38*(1), 149–161.

Geniole, S. N., Keyes, A. E., Carré, J. M., & McCormick, C. M. (2014). Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Personality and Individual Differences*, *57*, 59–64.

Gheorghiu, A. I., Callan, M. J., & Skylark, W. J. (2017). Facial appearance affects science communication.

Proceedings of the National Academy of Sciences, *114*(23), 5970–5975.

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: Facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 571–576.

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially affect the consistency of social evaluations. *Personality and Social Psychology Bulletin*, *41*(8), 1123–1134.

Hsieh, T.-S., Kim, J.-B., Wang, R. R., & Wang, Z. (2020). Seeing is believing? Executives' facial trustworthiness, auditor tenure, and audit fees. *Journal of Accounting and Economics*, *69*(1), 101260.

Hughes, B. L., Ambady, N., & Zaki, J. (2017). Trusting outgroup, but not ingroup members, requires control: Neural and behavioral evidence. *Social Cognitive and Affective Neuroscience*, *12*(3), 372–381.

Jaegera, B., Todorov, A. T., Evans, A. M., & Beest, I. van. (2020). Can we reduce facial biases? Persistent effects of facial trustworthiness on T sentencing decisions. *Journal of Experimental Social Psychology*, *90*, 104004.

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., …Coles, N. A. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour*, *5*, 159–169.

Kaisler, R. E., & Leder, H. (2016). Trusting the looks of others: Gaze effects of faces in social settings. *Perception*, *45*(8), 875–892.

Kaisler, R. E., & Leder, H. (2017). Combined effects of gaze and orientation of faces on person judgments in social situations. *Frontiers in Psychology*, *8*, 259.

Kleider-Offutt, H. M., Bond, A. D., & Hegerty, S. E. A. (2017). Black stereotypical features: When a face type can get you in trouble. *Current Directions in Psychological Science*, *26*(1), 28–33.

Kleisner, K., Priplatova, L., Frost, P., & Flegr, J. (2013). Trustworthy-looking face meets brown eyes. *PLoS ONE*, *8*(1), e53285.

Kosinski, M. (2017). Facial width-to-height ratio does not predict self-reported behavioral tendencies. *Psychological Science*, *28*(11), 1675–1682.

Krys, K., Hansen, K., Xing, C., Szarota, P., & Yang, M. (2014). Do only fools smile at strangers? Cultural differences in social perception of intelligence of smiling individuals. *Journal of Cross-Cultural Psychology*, *45*(2), 314–321.

Li, T., Liu, X., Pan, J., & Zhou, G. (2017). The interactive effect of facial appearance and behavior statement on trust belief and trust behavior. *Personality and Individual Difference*, *117*, 60-65.

Linke, L., Saribay, S. A., & Kleisner, K. (2016). Perceived trustworthiness is associated with position in a corporate hierarchy. *Personality and Individual Differences*, *99*, 22–27.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135.

Manssuer, L., Roberts, M., & Tipper, S. P. (2015). The late positive potential indexes a role for emotion during learning of trust from eye-gaze cues. *Social Neuroscience*, *10*(6), 635–650.

Mende-Siedlecki, P. (2018). Changing our minds: The neural bases of dynamic impression updating. *Current Opinion in Psychology*, *24*, 72–76.

Mende-Siedlecki, P., Cai, Y., & Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, *8*(6), 623–631.

Mende-Siedlecki, P., & Todorov, A. (2016). Neural dissociations between meaningful and mere inconsistency in impression updating. *Social Cognitive and Affective Neuroscience*, *11*(9), 1489–1500.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: Social-cognitive processing in the medial prefrontal cortex. *NeuroImage*, *26*(1), 251–257.

Neys, W. D. (2015). Adolescents gradually improve at detecting trustworthiness from the facial features of unknown

adults. *Journal of Economic Psychology*, *47*, 17–22.

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*(11), 566–570.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092.

Oosterhof, N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*, 128–133.

Paetzel, M., & Castellano, G. (2019). Let me get to know you better: Can interactions help to overcome uncanny feelings? In *Proceedings of the 7th International Conference on Human-Agent Interaction* (pp. 59–67). New York, USA: Association for Computing Machinery.

Qi, Y., Li, Q., & Du, F. (2018). Are rich people perceived as more trustworthy? Perceived socioeconomic status modulates judgments of trustworthiness and trust behavior based on facial appearance. *Frontiers in Psychology*, *9*, 512.

Quinn, P. C., Lee, K., & Pascalis, O. (2019). Face processing in infancy and beyond: The case of social categories. *Annual Review of Psychology*, *70*(1), 165–189.

Rescorla, R. A. (1976). Stimulus generalization: Some predictions from a model of Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(1), 88–96.

Rescorla, R. A., & Furrow, D. R. (1977). Stimulus similarity as a determinant of Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*(3), 203–215.

Rezlescu, C., Duchaine, B., Olivola, C. Y., & Chater, N. (2012). Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS ONE*, *7*(3), e34293.

Shen, X., Mann, T. C., & Ferguson, M. J. (2020). Beware a dishonest face?: Updating face-based implicit impressions using diagnostic behavioral information. *Journal of Experimental Social Psychology*, *86*, 103888.

Slepian, M. L., & Carr, E. W. (2019). Facial expressions of authenticity: Emotion variability increases judgments of trustworthiness and leadership. *Cognition*, *183*, 82–98.

Sofer, C., Dotsch, R., Oikawa, M., Oikawa, H., Wigboldus, D. H. J., & Todorov, A. (2017). For your local eyes only: Culture-specific face typicality influences perceptions of trustworthiness. *Perception*, *46*(8), 914–928.

Sofer, C., Dotsch, R., Wigboldus, D. H. J., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science*, *26*(1), 39–47.

Song, Y. (2020). Building a ‘deeper’ trust: Mapping the facial anthropomorphic trustworthiness in social robot design through multidisciplinary approaches. *The Design Journal*, *23*(4), 639–649.

Song, Y., & Luximon, Y. (2020). Trust in AI agent: A systematic review of facial anthropomorphic trustworthiness for social robot design. *Sensors*, *20*(18), 5087.

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–354.

Strachan, J. W. A., Kirkham, A. J., Manssuer, L. R., Over, H., & Tipper, S. P. (2017). Incidental learning of trust from eye-gaze: Effects of race and facial trustworthiness. *Visual Cognition*, *25*(7–8), 802–814.

Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, *108*(2), 397–415.

Todorov, A. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545.

Tsankova, E., & Kappas, A. (2016). Facial skin smoothness as an indicator of perceived trustworthiness and related traits. *Perception*, *45*(4), 400–408.

Verosky, S. C., Porter, J., Martinez, J. E., & Todorov, A. (2018). Robust effects of affective person learning on evaluation of faces. *Journal of Personality and Social Psychology*, *114*(4), 516–528.

Verosky, S. C., & Todorov, A. (2010). Differential neural responses to faces physically similar to the self as a function of their valence. *NeuroImage*, *49*(2), 1690–1698.

Wang, D., Nair, K., Kouchaki, M., Zajac, E. J., & Zhao, X. (2019). A case of evolutionary mismatch? Why facial width-to-height ratio may not predict behavioral tendencies. *Psychological Science*, *30*(7), 1074–1081.

Willis, M. L., Palermo, R., & Burke, D. (2011). Judging approachability on the face of it: The influence of face and body expressions on the perception of approachability. *Emotion*, *11*(3), 514.

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*(8), 1325–1331.

Wilson, J., & Rule, N. (2017). Advances in understanding the detectability of trustworthiness from the face: Toward a taxonomy of a multifaceted construct. *Current Directions in Psychological Science*, *26*, 396–400.

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2018). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science*, *21*(2), e12537.

Yu, M., Saleem, M., & Gonzalez, C. (2014). Developing trust: First impressions and experience. *Journal of Economic Psychology*, *43*, 16–29.

Zebrowitz, L. A., Fellous, J., Mignault, A., Andreoletti, C., Soc, P., Rev, P., … Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, *7*(3), 194–215.

Zebrowitz, L. A., & Montepare, J. M. (2005). Appearance DOES matter. *Science (New York, N.Y.)*, *308*, 1565–1566.

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517.

Zhang, D., Lin, H., & Perrett, D. I. (2020). Apparent emotional expression explains the effects of head posture on perceived trustworthiness and dominance, but a measure of facial width does not. *Perception*, *49*(4), 422–438.

Zhang, J., Li, S., Zhang, J.-Y., Du, F., Qi, Y., & Liu, X. (2020). A literature review of the research on the uncanny valley. In P.-L. P. Rau (Ed.), *HCII 2020, Lecture notes in computer science: Vol. 12192: Cross-cultural design. user experience of products, services, and intelligent environments* (pp. 255–268). Springer, Cham.

Regulation of facial trustworthiness evaluation: The proposal and empirical verification of the experience transfer hypothesis

QI Yue[1,2], QIN Shaotian[1,2], WANG Kexin[1,2], CHEN Wenfeng[1,2]

[1]*The Department of Psychology, Renmin University of China, Beijing* 100872, *China*

[2]*The Laboratory of the Department of Psychology, Renmin University of China, Beijing* 100872, *China*

Abstract: The impression of trustworthiness based on facial appearance plays an important role in interpersonal trust and cooperative behavior. Previous research results have indicated that a variety

of additional information (e.g., facial features, context) could affect facial judgments by both bottom-up and top-down processes. However, the mechanism of the two adjustment factors is still unclear. Focusing on this issue, the current study proposes the “experience transfer hypothesis”, and explores the cognitive mechanisms, the boundary conditions of the experience transfer effect, and the underlying process of these two adjustments. This study will provide new ideas and empirical evidence to deepen our understanding of how interpersonal trust builds.

Keywords: trust, moderating effect, facial trustworthiness, impression formation